

Computer-adaptive test to measure community reintegration of Veterans

Linda Resnik, PT, PhD, OCS;^{1–2*} Feng Tian;³ Pengsheng Ni, MD;³ Alan Jette, PhD, PT^{3–4}

¹Providence Department of Veterans Affairs Medical Center, Providence, RI; ²Department of Community Health, Brown University, Providence, RI; ³Health & Disability Research Institute, and ⁴Department of Health Policy and Management, Boston University School of Public Health, Boston, MA

Abstract—The Community Reintegration of Injured Service Members (CRIS) measure consists of three scales measuring extent of, perceived limitations in, and satisfaction with community reintegration. Length of the CRIS may be a barrier to its widespread use. Using item response theory (IRT) and computer-adaptive test (CAT) methodologies, this study developed and evaluated a briefer community reintegration measure called the CRIS-CAT. Large item banks for each CRIS scale were constructed. A convenience sample of 517 Veterans responded to all items. Exploratory and confirmatory factor analyses (CFAs) were used to identify the dimensionality within each domain, and IRT methods were used to calibrate items. Accuracy and precision of CATs of different lengths were compared with the full-item bank, and data were examined for differential item functioning (DIF). CFAs supported unidimensionality of scales. Acceptable item fit statistics were found for final models. Accuracy of 10-, 15-, 20-, and variable-item CATs for all three scales was 0.88 or above. CAT precision increased with number of items administered and decreased at the upper ranges of each scale. Three items exhibited moderate DIF by sex. The CRIS-CAT demonstrated promising measurement properties and is recommended for use in community reintegration assessment.

Key words: community reintegration, computer-adaptive test, disability, factor analysis, measurement, military healthcare, outcomes assessment, participation, rehabilitation, Veterans.

INTRODUCTION

Over the past decade, more than 2 million U.S. servicemembers have deployed to Iraq and Afghanistan

(Operation Iraqi Freedom/Operation Enduring Freedom [OIF/OEF]). Studies of OIF/OEF Veterans report a high prevalence of problems related to posttraumatic stress disorder (PTSD), anxiety, major depression, and mild traumatic brain injury [1–4], which can pose substantial challenges to community reintegration. Helping our newest cohort of combat Veterans adjust to life at home and in the community and return to healthy participation in major social life roles is a priority.

The early identification and prevention of problems in the community reintegration of combat-deployed Veterans and the evaluation of clinical interventions to promote healthy social role functioning require accurate assessment and monitoring of community reintegration.

Abbreviations: CAPI = computer-assisted personal interviews, CAT = computer-adaptive test, CFA = confirmatory factor analysis, CFI = comparative fit index, CRIS = Community Reintegration of Injured Service Members, DIF = differential item functioning, EFA = exploratory factor analysis, ICF = International Classification of Health and Functioning, IRT = item response theory, OIF/OEF = Operation Iraqi Freedom/Operation Enduring Freedom, PTSD = posttraumatic stress disorder, PVAMC = Providence Department of Veterans Affairs Medical Center, RMSEA = root mean square error approximation, TLI = Tucker–Lewis Index, VA = Department of Veterans Affairs.

*Address all correspondence to Linda Resnik, PT, PhD, OCS; Research Health Scientist, Providence VA Medical Center, 830 Chalkstone Ave, Providence, RI 02908; 401-273-7100, ext 2368. Email: Linda.Resnik@va.gov

<http://dx.doi.org/10.1682/JRRD.2011.04.0081>

Although numerous instruments that measure aspects of community reintegration exist, they lack demonstrated validity for use with OIF/OEF Veterans. There is currently no accepted gold standard measure against which a Veteran community reintegration measure can be compared and no universal agreement on the most meaningful content areas for assessing community reintegration of Veterans. To address this gap, we developed a new Veteran-centric measure of community reintegration, the Community Reintegration of Injured Service Members (CRIS).

The CRIS measure was designed to assess the construct of participation as defined by the World Health Organization's International Classification of Health and Functioning (ICF) [5–6], which we consider synonymous with community reintegration. Using the domain of Participation as defined by the ICF approach to characterize community reintegration is consistent with recent recommendations of the Department of Veterans Affairs' (VA's) State of the Science Working Group on Community Reintegration [7]. The development work for the original version of the CRIS was conducted in a three-stage process designed to maximize content validity for use with Veterans. In the exploratory stage, dimensions of and challenges in community reintegration were identified through a qualitative study of injured servicemembers, caregivers, and clinicians. Veterans and caregivers discussed challenges in daily life; mobility; activities at home in the neighborhood and community; and family, social, and work life. Clinicians discussed the common challenges they had observed in their OIF/OEF patients. We used a directed approach to content analysis [8] to code the data by using the ICF domains of Activities and Participation. Because the CRIS was intended to measure the concept of participation, not activity as defined by the ICF, we coded all challenges relating to complex tasks and/or societal involvement. We included complex activities that might be undertaken alone and not with others, such as eating, thinking, and traveling, as aspects of role participation.

Items addressing each problem area identified in the qualitative data were generated to form the initial CRIS item pool, drawing from and adapting questions from existing measures based on our review of the literature. When necessary, we developed new items following two guidelines. First, for each problem, separate items were written for each of three dimensions: extent, perceived limitations, and satisfaction. Second, all questions were

phrased to facilitate comprehension by assessing current life situation, with no comparison to life before injury or to other persons without injury or who had not been deployed. In addition, questions were framed in the present or within the last 2 weeks to minimize recall bias.

The CRIS instrument was further revised in the confirmatory stage, during which a multidisciplinary group of clinicians (including primary care and polytrauma physicians, physical and occupational therapists, recreational therapist, and psychologists) reviewed the item set and provided comments on the wording, content, and importance. Finally, we conducted cognitive-based interviews with seven OIF/OEF Veterans that involved asking respondents to "think out loud" as they answered questions. Respondents were asked to talk about their response process, including their comprehension of the item, their ability to recall the answer, and their strategy of retrieving information related to the question [9]. Further refinements in the wording of CRIS items were made based on cognitive testing.

Following these revisions, we conducted two pilot studies to examine unidimensionality, internal consistency, reliability, concurrent validity, and construct validity. Following the first pilot study, we revised all misfit items and added items to address the higher and lower ends of participation. We performed cognitive-based testing on new items with a sample of six Veterans recruited from the Providence VA Medical Center (PVAMC). We then conducted a second pilot study of the psychometrics of the CRIS. These analyses found that the scales constructed from the original item sets were unidimensional, with Rasch models predicting the majority of variance in the data for each scale (95-item Extent = 0.53, 107-item Perceived Limitations = 85.2, 85-item Satisfaction = 73.3).

The CRIS fixed-form measure was developed from the full CRIS scale item sets to provide a briefer alternative to administration of the entire CRIS item set. The fixed-form CRIS includes 151 items within three separate subscales. Each subscale contains items from the 9 Activity and Participation content domains as defined by the ICF and includes items related to negative as well as positive aspects of community reintegration [10]. In developing the fixed-form measure, we included items that had good reliability and represented important content areas identified in formative research as well as a broad spectrum of item difficulty as calculated by the preliminary item response theory (IRT) analysis. Preliminary analyses confirmed the concurrent and construct validity of the

fixed-form scales. Additional details regarding the development and pilot testing of the original CRIS and CRIS fixed-form measure can be found elsewhere [10].

We believe that the CRIS measure is unique as compared with other community reintegration scales in that it is based on the conceptual foundation of the ICF and contains separate scales, each measuring a different aspect of community reintegration. The CRIS scales can be used singly or in combination, chosen to best match the purpose of data collection. The CRIS Extent scale measures extent/frequency, which is an important aspect of assessment of community reintegration. However, assessment of frequency and amount of activity does not consider individual preferences, personal choices, life stages, and values. The assessment of subjective aspects, such as perceived limitations and satisfaction, add depth to the assessment of community reintegration.

The original CRIS scales are lengthy, and the three fixed-form scales developed from the item set, though briefer, take approximately 30 minutes to administer [10], rendering its administration impractical in busy clinical and research settings. Thus, the purpose of this study was to utilize contemporary measurement methods to develop a computer-adaptive test (CAT) version of the CRIS, the CRIS-CAT, that would allow accurate and precise measurement of community reintegration with reduced respondent burden. The goal is a briefer outcome instrument that would be an alternative to the fixed-form CRIS for use by research and policy makers, as well as a valuable addition to the clinical information collected for and available to clinicians.

METHODS

Selection and Description of Participants

Subjects were English-speaking men and women, ages 18–59, stratified by expected level of community reintegration to ensure a range of degree of community reintegration in the sample. Group 1 subjects were non-OIF/OEF Veterans with stable housing, steady employment, and negative screens for depression, PTSD, and substance abuse history; group 2 subjects were non-OIF/OEF Veterans who were homeless or at risk for homelessness because of insecure housing and/or chronic unemployment; and group 3 subjects were Veterans of the OEF or OIF conflicts. Subjects were recruited from the primary care and PTSD services of the PVAMC;

through community and military reserve sites in Rhode Island, southeastern Massachusetts, and nearby Connecticut; and through homeless shelters in the region.

Data Collection

The CRIS-CAT items were administered by trained interviewers using computer-assisted personal interviews (CAPI). The CAPI system had standard protections against entering out-of-range values and allowed direct data downloading into a central database. Demographic information (age, sex, ethnicity, race, education, living status, and housing status) was collected for each subject. After obtaining informed consent from each subject, the interviewer read the standardized CRIS instructions from a script and showed the subject the choice of response categories. Subjects indicated their chosen responses, which were documented by the interviewer.

DATA ANALYSIS

Factor Analysis

Both IRT and CAT systems assume that all items in a scale measure a single, unitary concept, often referred to as assumptions of unidimensionality and local independence [11–13]. Item sets that violate these assumptions may be less effective in achieving appropriate modeling of the data and may limit the accuracy of a CAT instrument. The dimensionality of responses to the CRIS item pool within each scale was evaluated using exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). Because the items were polytomous, we used a weighted least squares method with mean and variance correction estimator based on a polychoric correlation matrix, which is more precise when analyzing moderate-sized samples with skewed categorical data [14–15]. Three initial EFAs and three initial unidimensional CFAs were conducted with items in each of the three CRIS scales, respectively: Extent ($N = 123$ items), Perceived Limitations ($N = 172$ items) and Satisfaction ($N = 124$ items). Example items in each CRIS scale are displayed in **Table 1**.

In the EFA analysis, the factor loadings, eigenvalues, and percentage of variance explained by the first factor were used to assess unidimensionality. In the unidimensional CFA analysis, the model fit was assessed by multiple fit indexes, including comparative fit index (CFI), Tucker–Lewis Index (TLI), root mean square error approximation (RMSEA), and residual correlations. CFI and TLI values

Table 1.
Sample items in each Community Reintegration of Injured Service Members scale.

Scale	Item Content	Response Set
Extent	How often did you get enough sleep?	7-point Frequency
	How often did you get together with friends?	7-point Frequency
	How often did you do your chores where you lived?	7-point Frequency
	How often did you eat fruits or vegetables?	7-point Frequency
	How often did you follow current events?	7-point Frequency
Perceived Limitations	I avoided driving a car.	7-point Agreement
	I did not have a place to live that met my needs.	7-point Agreement
	I did not have enough money to cover my bills.	7-point Agreement
	I felt I easily lost control of my feelings.	7-point Agreement
	I got along with my children.	7-point Agreement
Satisfaction	How satisfied were you with getting enough to eat?	7-point Happiness
	How satisfied were you with your ability to concentrate on what you were doing?	7-point Happiness
	How satisfied were you with your daily accomplishments?	7-point Happiness
	How satisfied were you with the way you protected yourself from harm?	7-point Happiness
	How satisfied were you with your ability to prepare meals?	7-point Happiness

range from 0 to 1 and values of 0.90 or higher suggest acceptable fit; RMSEA values less than 0.08 mean acceptable fit. Both EFA and CFA were conducted using the MPlus software (Muthen & Muthen; Los Angeles, California) [16]. Local independence was evaluated by inspecting the residual correlations between items, also using MPlus software [16]. Items with residual correlations greater than ± 0.2 were considered to show local dependence [17–19].

Item Response Theory Analysis

Given our relatively small sample size for IRT analyses, the one-parameter Rasch partial-credit model [20] was used as the IRT-based methodology for all three scales. In the application of IRT models, the item parameters have to be calculated or estimated, a procedure referred to as calibration. The item parameters and fit statistics in this study were calculated for each scale using WINSTEPS (Chicago, Illinois) [21], which is based on joint maximum likelihood estimation. The item fit statistics for each item are based on the comparison of expected and observed values. The infit mean square (also referred to as the weighted mean square) is derived from the squared standardized residuals for each item/person interaction. The infit mean squares are weighted by their individual variance to minimize the impact of unexpected responses far from the measure. Items demonstrating more variation than predicted by the model can be considered as not conforming to the unidimensionality requirement of the Rasch model. Values between 0.5 and 1.5 indicate productive measurement

[21]. To evaluate breadth of coverage in each scale, we calculated item category parameter distributions for each item bank compared with the sample distribution.

The item responses should only depend on the subject's ability level and the statistical characteristics of the item. Significant differential item functioning (DIF) indicates that variables other than the subject's ability level are influencing the item response. We used ordinal logistic regression models to assess DIF across age, sex, education, and race. Based on Jodoin and Gierl's approach [22], we used R-square change to classify DIF. An R-square change less than 0.035 indicated no DIF, a value between 0.035 and 0.07 indicated moderate DIF, and a value greater than 0.07 indicated large DIF.

Computer-Adaptive Test Analysis

We examined the accuracy of CAT algorithms for the three CRIS-CAT scales (Extent, Perceived Limitations, and Satisfaction) by using real-data computer simulation methods. The simulation program, developed at Boston University, includes several item-selection methods, the option to set content balancing or not (in this simulation, we set the first nine items to be content balanced across subcontent domains), different score estimation algorithms (in this simulation, the weighted likelihood estimation was used), and several stopping rules (such as fixed minimum and maximum number of items, the level of score precision or both). We used a real-data simulation approach for investigating the merits of CAT. The scales score estimates were based on the complete set of all actual item responses in

each domain (called the IRT criterion score); this served as the criterion standard against which scores from the CRIS-CAT were compared. As items were selected for administration in the simulation, responses were taken from the actual data set. After each response, we estimated a score based on all administered items to that point in the simulation and calculated the associated standard error. The selection of the next item was based on the item that could provide the most information at the estimated score. We established specific stop-rules based on the number of items in the CAT (5, 10, 15, or 20 items). We also set the stopping as a minimum of 10 items, maximum of 20 items, and reliability of 0.9. These simulated scores were compared with the criterion standard using intraclass correlation coefficient (3,1).

To evaluate the precision of the CAT score, we calculated and compared standard errors associated with each subject's score for the 5-, 10-, 15-, and 20-item CATs and the varied number of items CAT with full item bank.

RESULTS

The sample consisted of 517 Veterans: 69 in group 1 (those with better community reintegration), 99 in group 2 (homeless Veterans), and 332 in group 3 (OIF/OEF Veterans). An additional 17 Veterans were screened into one of the above groups but, on inspection of their data, were found not to meet inclusion criteria for any of these groups. The characteristics of subjects in our 517-person sample are shown in **Table 2**.

In the initial EFA run for the Extent, Satisfaction, and Perceived Limitations scales, 9 out of the 123 items in the Extent scale, 1 out of the 124 items in the Satisfaction scale, and 2 out of the 172 items in the Perceived Limitations scale had very low factor loadings of less than 0.2 when only one factor was extracted. These items were removed from further analysis. In the initial CFA run for the scales, 37 items in the Extent scale, 35 items in the Satisfaction scale, and 20 items in the Perceived Limitations scale showed local dependence with residual correlations greater than 0.2. These items were also removed from further analysis.

In the EFA of the remaining items of the three scales (77 items in Extent scale, 88 items in Satisfaction scale, and 150 items in Perceived Limitations scale), the first factor explained from 33 percent to about 53 percent of the total variance and the ratio between the eigenvalue of the

Table 2.
Characteristics of sample ($N = 517$).

Characteristic	Mean \pm SD, Range
CRIS-CAT Scale	
Extent	49.3 \pm 9.4, 24–90
Perceived Limitations	49.7 \pm 9.4, 33–99
Satisfaction	49.7 \pm 9.3, 34–91
Age	39.7 \pm 11.7, 19–60
Time Since Return from Deployment (mo)	74.8 \pm 120.8, 0–480
	Frequency (%)
Sex	
Male	433 (83.8)
Female	84 (16.3)
Race/Ethnicity	
White	394 (76.5)
Black	44 (8.5)
Other	44 (8.5)
Mixed	33 (6.4)
Hispanic	43 (8.4)
Has Children	317 (61.3)
Lives with Children Under Age 18	185 (36.2)
Education	
Less than High School	7 (1.4)
High School	121 (23.4)
GED	30 (5.8)
Some College	221 (42.8)
College	99 (19.2)
Postgraduate	39 (7.5)
Employment Status	
Unemployed	99 (19.2)
Not Working Because of Disability/Medical Hold	87 (16.9)
Working Part-Time/Training	56 (10.9)
Working Full-Time	270 (52.4)
Retired	3 (0.6)
Income (\$)	
<25,000	173 (33.7)
25,000–50,000	146 (28.4)
>50,000	195 (37.9)
Marital Status	
Unmarried	181 (35.0)
Married	212 (41.0)
Divorced, Separated, or Widowed	124 (24.0)
Residence	
Outside	2 (0.4)
Staying with Friend	47 (9.1)
Veteran Home	31 (6.0)
House	144 (27.9)
Apartment	272 (52.6)
Other	21 (4.1)
Depression Diagnosis (Yes)	154 (30.3)
PTSD Diagnosis (Yes)	144 (28.4)
Mental Illness Diagnosis (Yes)	105 (20.6)
Alcohol/Drug Abuse Diagnosis (Yes)	141 (27.3)
Ever Deployed to Combat Area (Yes)	354 (68.5)
OIF/OEF (Yes)	332 (64.2)

CRIS-CAT = Community Reintegration of Injured Service Members Computer Adaptive Test, GED = general equivalency diploma, OIF/OEF = Operation Iraqi Freedom/Operation Enduring Freedom, PTSD = posttraumatic stress disorder, SD = standard deviation.

first factor and the eigenvalue of the second factor ranged from about 9 to about 12. The CFA of remaining items in the three scales also showed acceptable model fit: CFI = 0.905, TFI = 0.903, and RMSEA = 0.048 for the Extent scale; CFI = 0.916, TFI = 0.914, and RMSEA = 0.066 for the Satisfaction scale; and CFI = 0.906, TFI = 0.905, and RMSEA = 0.053 for the Perceived Limitations scale. The factor analysis results are summarized in **Table 3**.

Acceptable item fit statistics were found for the three scales in the IRT analysis. Some misfit was demonstrated by 2 out of the 77 Extent items, 5 out of the 88 Satisfaction items, and 9 out of the 150 Perceived Limitations items. However, the 2 misfitting Extent items, 3 of the 5 misfitting Satisfaction items, and 3 of the 9 misfitting Perceived Limitations items were kept after expert review because of the importance of their contents. So the final item pool contained 77 items in the Extent scale, 86 items in the Satisfaction scale, and 144 items in the Perceived Limitations scale. The calibration results are summarized in the **Appendix** (available online only).

No items showed DIF in the age, education, or race variables across the three scales. For sex, two items (How often did you spend quality time with your children? and How often did you avoid going out alone after dark?) showed moderate DIF in the Extent scale and one item (I

avoided going out alone after dark.) showed moderate DIF in the Perceived Limitations scale.

Table 4 displays the accuracy correlations for CATs of variable sizes with the overall item pools for each content scale of the CRIS-CAT instrument. Accuracy of the variable-item CATs for all three scales was 0.88 or above. For the variable size CAT, the accuracy coefficient was 0.93 for the Extent scale and 0.95 for the other two scales.

Figure ((a)–(c)) displays the precision of the CRIS-CATs of different sizes in units of standard error of the measure as compared with the item pools. As expected, the precision of the item pools exceeded that of the CATs and the CAT precision increased with the number of items administered. Precision decreased at the upper ranges of the continuum of each CRIS-CAT content scale.

We examined the distribution of items in each content item pool to see how it matched the score distribution in the study sample. For all three content scales, a large degree of overlap existed between the item category distribution and the sample score distribution. In the Satisfaction scale, few items were at the higher end of the continuum where there were some scores in this sample. A better match existed between sample and item distributions in the other two scales.

Table 3.

Summary statistics for factor analysis results for each Community Reintegration of Injured Service Members (CRIS) scale.

CRIS Scale	CFA			EFA	
	TLI	CFI	RMSEA	% Total Variance by 1st Factor	Eigenvalue Ratio Between 1st & 2nd Factors
Extent	0.905	0.903	0.048	33	9
Satisfaction	0.916	0.914	0.066	53	15
Perceived Limitations	0.906	0.905	0.053	50	11.6

CFA = confirmatory factor analysis, CFI = comparative fit index, EFA = exploratory factor analysis, RMSEA = root mean square error approximation, TLI = Tucker–Lewis Index.

Table 4.

Accuracy (confidence interval) of different size computer-adaptive tests (CATs) for each Community Reintegration of Injured Service Members (CRIS) scale.

CRIS Scale	5-CAT	10-CAT	15-CAT	20-CAT	Variable No. of Items in CAT*
Extent	0.773 (0.737,0.806)	0.882 (0.862,0.900)	0.918 (0.903,0.930)	0.947 (0.937,0.955)	0.926 (0.913,0.937)
Perceived Limitations	0.871 (0.848,0.890)	0.929 (0.916,0.940)	0.953 (0.945,0.960)	0.967 (0.961,0.972)	0.948 (0.939,0.978)
Satisfaction	0.897 (0.879,0.913)	0.932 (0.919,0.942)	0.958 (0.950,0.964)	0.971 (0.966,0.976)	0.949 (0.940,0.957)

*Minimum number of items = 10, maximum = 20, reliability \geq 0.9.

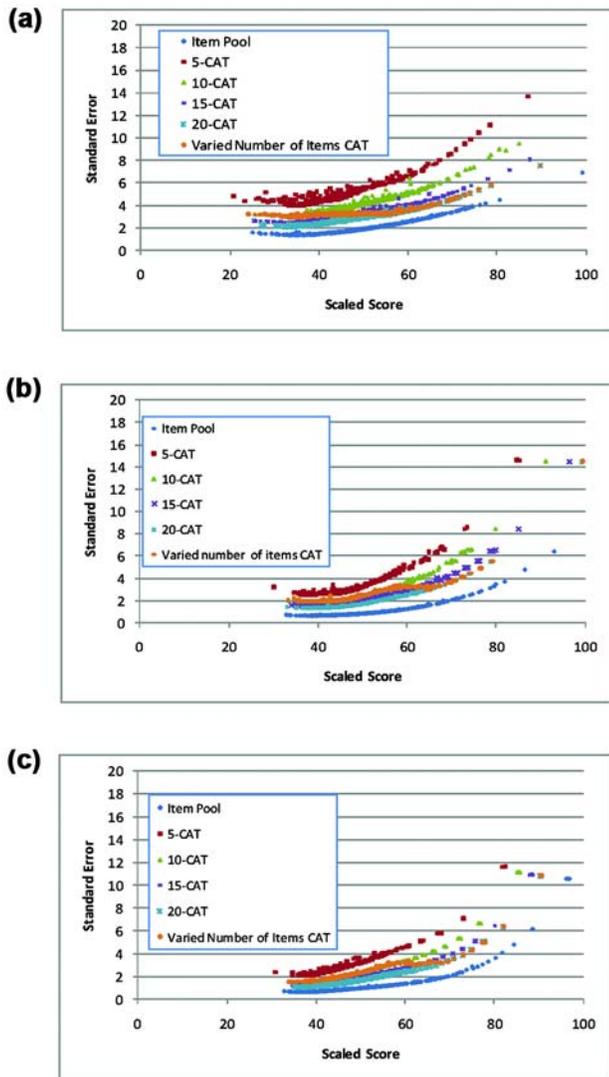


Figure.

Precision of different size computer-adaptive tests (CATs) for Community Reintegration of Injured Service Members scales: (a) Extent, (b) Perceived Limitations, and (c) Satisfaction.

DISCUSSION

Caring for complex patients within the limitations of brief treatment visits is a challenge for busy providers because acute symptoms and concerns tend to “crowd out” less urgent needs. Presently, neither the VA’s nor the Department of Defense’s electronic medical record contains standardized data elements related to community reintegration. Monitoring Veteran community reintegration following deployment is critically important for pub-

lic health. To date, only one validated instrument, the CRIS, has been developed to measure Veteran community reintegration. However, its length may diminish the feasibility of its widespread implementation. To overcome this limitation, we used contemporary measurement methods to develop a CAT version of the CRIS, called the CRIS-CAT.

The results of our analyses confirmed three distinct unidimensional domains of community reintegration (Extent, Perceived Limitations, and Satisfaction) included in the CRIS and revealed that all three CRIS-CAT scales performed well in our sample of Veterans. The results of this study indicate that the 10-item CATs for Perceived Limitations and Satisfaction and the 20-item CAT for Extent will maximize psychometric properties with minimal time (estimated 10 minutes) needed for data collection. Use of the CRIS-CAT instead of a fixed-form measure of community reintegration will reduce barriers to routine data collection, making collection of this outcome measure more feasible to implement.

We observed a large number of items in the lower end of the continuum of the Extent scale as compared with the number of low scores in our sample. This may be due to the nature of our sample, most of which was recruited from outpatient healthcare services and/or community locations and workplaces. Thus, we expected that overall our sample would evidence higher scores in the frequency of their participation in activities as compared with those who were more severely injured as a result of polytraumatic injuries, homebound, or receiving inpatient care. We attempted to capture the lower range of community reintegration scores by including a subsample of persons who were homeless or at-risk of becoming homeless because of chronic unemployment. This group of Veterans did have lower scores on all CRIS-CAT scales. Further research is needed to examine the item-person fit for the Extent scale in a sample that is expected to be more severely impacted. We also observed few items at the higher end of the Satisfaction scale, suggesting a need to develop additional items at the higher end, to minimize potential ceiling effects.

Our analyses revealed moderate DIF by sex in three items. These items related to going out alone after dark and spending quality time with children. These items might function differently by sex. However, our finding needs to be verified using another data set. If this finding persists, we might consider calibrating those items separately by sex. Because of limitations of our current sample

size, we are unable to do this and will continue to monitor these items in the CAT performance.

We recognize that clinicians and researchers may wish to compare scores of the CRIS-CAT with scores from the original fixed-form CRIS instrument. Because the CRIS-CAT summary score is the appropriate statistic generated by the Rasch model, we used the weighted likelihood estimation method to create the equivalent CRIS fixed-form summary score based on the CRIS-CAT scores. These equivalent scores are roughly comparable, though not identical, and the calculations are available from the authors by request.

When we examined the comparability of items in the CRIS-CAT scales and the CRIS fixed-form scales, we found that 10 items in the CRIS fixed-form Extent scale, 6 items in the Perceived Limitations scale, and 5 items in the Satisfaction scale were omitted from the CRIS-CAT scale item pool. This loss of items may be attributable to the differences in data analyzed for each study. The preliminary studies' small sample sizes resulted in missing data on numerous items (particularly those related to parenting and employment). Thus, many items from the original CRIS item pool were not included in the preliminary Rasch models. The current study involved a much larger sample and employed a more robust sampling strategy, recruiting a more heterogeneous population that included Veterans from the community (rather than from a single medical center). Therefore, the final Rasch models that we developed for the CRIS-CAT scales included more items from the CRIS item pool (not simply those in the scales identified in preliminary research). Thus, it is not surprising that some differences existed in item content of the CRIS scales identified in this research and the preliminary research. These differences suggest that direct comparison of CRIS fixed-form and CRIS-CAT scores should be made using our transformations.

One limitation of our analyses is that our conclusions are based on real-data simulations, which assume that respondents would answer the subset of items selected using CAT in an identical manner to the answers they would provide when responding to those items embedded in the larger item site. Though they are considered good approximations, data simulations like these are not perfect simulations of actual CAT administration and may overestimate these correlations [23]. Future research is needed to examine the accuracy of the CRIS-CAT and assess the administrative burden of implementing the CRIS-CAT in prospective studies.

Further research is needed to examine the CRIS-CAT utility in non-Veterans with mental health conditions and those who have experienced physical and psychological trauma.

CONCLUSIONS

We used the CRIS item pool to develop and evaluate a brief community reintegration measure by using IRT and CAT methodologies. The CRIS-CAT demonstrated promising measurement properties. Acceptable item fit statistics were found for final models. Accuracy of 10-, 15-, 20-, and variable-item CATs for all three scales was 0.88 or above. CAT precision increased with the number of items administered and decreased at the upper ranges of each scale. Three items exhibited moderate DIF by sex.

The strong conceptual basis for this CRIS-CAT scale combined with its ability to produce precise estimates with minimal respondent burden (approximately 10 minutes) make it a strong candidate for measurement of community reintegration of Veterans and a briefer alternative to the CRIS fixed-form measure.

We believe that routine assessment of community reintegration would enhance patient assessment and targeting of referrals to services such as mental health, social services, and Veterans Health Administration benefit programs, as well as inform primary care and other interventions that address underlying factors related to poor community reintegration. The CRIS-CAT would be an appropriate outcome measure for use by rehabilitation disciplines treating patients with polytraumatic injury.

ACKNOWLEDGMENTS

Author Contributions:

Study design: P. Ni, A. Jette.

Preparation of data: L. Resnik.

Interpretation of study results: L. Resnik, A. Jette.

Factor analyses: F. Tian, P. Ni.

IRT analyses and interpretation: P. Ni, A. Jette.

Other data analytic activities: F. Tian.

Manuscript writing: L. Resnik, F. Tian, P. Ni, A. Jette.

Obtained funding: L. Resnik.

Scientific leadership: L. Resnik, A. Jette (IRT analyses).

Financial Disclosures: The authors have declared that no competing interests exist.

Funding/Support: This material was based on work supported by the VA Health Services Research and Development Service (grant DHI-07-144).

Additional Contributions: The authors wish to acknowledge the analytical assistance of Matthew Borgia and the recruitment and data collection efforts of Regina Lynch, Pam Steager, and Julie Lyons.

Institutional Review: The institutional review board at the PVAMC approved the study, and all participants provided informed consent.

Participant Follow-Up: The authors do not plan to inform participants of the publication of this study.

Disclaimer: The information in this article does not necessarily reflect the position or policy of the Government; no official endorsement should be inferred.

REFERENCES

- Hoge CW, Auchterlonie JL, Milliken CS. Mental health problems, use of mental health services, and attrition from military service after returning from deployment to Iraq or Afghanistan. *JAMA*. 2006;295(9):1023–32. [\[PMID:16507803\]](#)
<http://dx.doi.org/10.1001/jama.295.9.1023>
- Milliken CS, Auchterlonie JL, Hoge CW. Longitudinal assessment of mental health problems among active and reserve component soldiers returning from the Iraq war. *JAMA*. 2007;298(18):2141–48. [\[PMID:18000197\]](#)
<http://dx.doi.org/10.1001/jama.298.18.2141>
- Institute of Medicine. *Returning home from Iraq and Afghanistan: Preliminary assessment of readjustment needs of veterans, service members, and their families*. Washington (DC): The National Academies Press; 2010.
- Sayer NA, Noorbaloochi S, Frazier P, Carlson K, Gravely A, Murdoch M. Reintegration problems and treatment interests among Iraq and Afghanistan combat veterans receiving VA medical care. *Psychiatr Serv*. 2010;61(6):589–97. [\[PMID:20513682\]](#)
<http://dx.doi.org/10.1176/appi.ps.61.6.589>
- Ustün TB, Chatterji S, Bickenbach J, Kostanjsek N, Schneider M. *The International Classification of Functioning, Disability and Health: A new tool for understanding disability and health*. *Disabil Rehabil*. 2003;25(11–12):565–71. [\[PMID:12959329\]](#)
<http://dx.doi.org/10.1080/0963828031000137063>
- World Health Organization. *International Classification of Functioning, Disability, and Health*. Geneva (Switzerland): World Health Organization; 2001.
- Resnik L, Bradford D, Glynn S, Jette A, Johnson Hernandez C, Wills S. Issues in defining and measuring veteran community reintegration: Proceedings of the Community Reintegration Working Group, VA Rehabilitation Outcomes Conference, Miami, Florida. *J Rehabil Res Dev*. 2012;49(1):87–100. [\[PMID:22000000\]](#)
<http://dx.doi.org/10.1682/JRRD.2010.06.0107>
- Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. 2005;15(9):1277–88. [\[PMID:16204405\]](#)
<http://dx.doi.org/10.1177/1049732305276687>
- Jobe JB, Mingay DJ. Cognitive research improves questionnaires. *Am J Public Health*. 1989;79(8):1053–55. [\[PMID:2751028\]](#)
<http://dx.doi.org/10.2105/AJPH.79.8.1053>
- Resnik L, Plow M, Jette A. Development of CRIS: measure of community reintegration of injured service members. *J Rehabil Res Dev*. 2009;46(4):469–80. [\[PMID:19882482\]](#)
<http://dx.doi.org/10.1682/JRRD.2008.07.0082>
- Thissen D, Reeve BB, Bjorner JB, Chang CH. Methodological issues for building item banks and computerized adaptive scales. *Qual Life Res*. 2007;16(Suppl 1):109–19. [\[PMID:17294284\]](#)
<http://dx.doi.org/10.1007/s11136-007-9169-5>
- Humbleton R. Applications of item response theory to improve health outcomes assessment: Developing item banks, linking instruments, and computer adaptive testing. In: Lipscomb J, Gotay C, Snyder C, editors. *Outcomes assessment in cancer*. Cambridge (UK): Cambridge University Press; 2005. p. 445–64.
- Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA, Thissen D, Revicki DA, Weiss DJ, Hambleton RK, Liu H, Gershon R, Reise SP, Lai JS, Cella D; PROMIS Cooperative Group. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*. 2007;45(5 Suppl 1):S22–31. [\[PMID:17443115\]](#)
<http://dx.doi.org/10.1097/01.mlr.0000250483.85507.04>
- Mislevy R. Recent developments in the factor analysis of categorical variables. *J Educ Behav Stat*. 1986;11:3–31. [\[PMID:34000000\]](#)
<http://dx.doi.org/10.3102/10769986011001003>
- Beauducel A, Herzberg P. On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Struct Equ Modeling*. 2006;13:186–203. [\[PMID:16500000\]](#)
http://dx.doi.org/10.1207/s15328007sem1302_2
- Muthen L, Muthen B. *MPlus statistical analysis with latent variables user's guide*. Los Angeles (CA): Muthen & Muthen; 2007.
- Yen W. Scaling performance assessments: strategies for managing local item dependence. *J Educ Meas*. 1993;30(3):187–213. [\[PMID:12300000\]](#)
<http://dx.doi.org/10.1111/j.1745-3984.1993.tb00423.x>
- Tate R. A comparison of selected empirical methods for assessing the structure of responses to test items. *Appl Psychol Meas*. 2003;27(3):159–203. [\[PMID:12300000\]](#)
<http://dx.doi.org/10.1177/0146621603027003001>

19. Chen WH, Thissen D. Local dependence indexes for item pairs using item response theory. *J Educ Behav Stat.* 1997; 22:265–89.
20. Masters GN. A Rasch model for partial credit scoring. *Psychometrika.* 1982;47(2):149–74.
<http://dx.doi.org/10.1007/BF02296272>
21. Linacre JM. A user's guide to WINSTEPS. Chicago (IL): MESA Press; 1999.
22. Jodoin MG, Gierl MJ. Evaluating Type I error and power rates using an effect size measure with the logistic regression procedure for DIF detection. *Appl Meas Educ.* 2001; 14:329–49.
http://dx.doi.org/10.1207/S15324818AME1404_2
23. Jette AM, McDonough CM, Ni P, Haley SM, Hambleton RK, Olarsch S, Hunter DJ, Kim YJ, Felson DT. A functional difficulty and functional pain instrument for hip and knee osteoarthritis. *Arthritis Res Ther.* 2009;11(4):R107.

[PMID:19589168]

<http://dx.doi.org/10.1186/ar2760>

Submitted for publication April 29, 2011. Accepted in revised form October 21, 2011.

This article and any supplementary material should be cited as follows:

Resnik L, Tian F, Ni P, Jette A. Computer-adaptive test to measure community reintegration of Veterans. *J Rehabil Res Dev.* 2012;49(4):557–66.

<http://dx.doi.org/10.1682/JRRD.2011.04.0081>

